Tools for Data Gathering and Handling & Best Practices

National FOT sim$^{TD}$

April 25th, 2013
Horst Rechner
Fraunhofer FOKUS
National FOT sim\textsuperscript{TD}: Project Introduction

Facts and figures

- National FOT in the Frankfurt / Hesse area
- 120 vehicles + 3 motorcycles
- Pool of 500 drivers with 8 week rotation
- FOT duration 6 months
- 1.6 Mio. Kilometers and 42000 hours of driving
- 85 test scenarios with multiple runs per scenario

Special in respect to data handling

- Separate team for test execution leading to…
- Exact scenario definition
- High degree of automation and abstraction
- Distributed environment (VIS, RIS, CIS)
# Consortium

<table>
<thead>
<tr>
<th>simTD: Partners</th>
<th>simTD: Supporters</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Vehicle Manufacturers</strong></td>
<td><strong>Federal Ministries</strong></td>
</tr>
<tr>
<td>Audi</td>
<td>Bundesministerium für Wirtschaft und Technologie</td>
</tr>
<tr>
<td>BMW</td>
<td>Bundesministerium für Bildung und Forschung</td>
</tr>
<tr>
<td>DAIMLER</td>
<td>Bundesministerium für Verkehr, Bau und Stadtentwicklung</td>
</tr>
<tr>
<td>Ford</td>
<td></td>
</tr>
<tr>
<td>Opel</td>
<td></td>
</tr>
<tr>
<td>VOLKSWAGEN</td>
<td></td>
</tr>
<tr>
<td></td>
<td>HESSEN</td>
</tr>
<tr>
<td>Netzbetreiber</td>
<td>Bundesministerium für Wirtschaft und Technologie</td>
</tr>
<tr>
<td>Deutsche Telekom</td>
<td>Bundesministerium für Bildung und Forschung</td>
</tr>
<tr>
<td></td>
<td>Bundesministerium für Verkehr, Bau und Stadtentwicklung</td>
</tr>
<tr>
<td>BMW</td>
<td>Hessen Mobil Straßen- und Verkehrsmanagement</td>
</tr>
<tr>
<td>TUM</td>
<td>Stadt Frankfurt am Main</td>
</tr>
<tr>
<td>HTW</td>
<td></td>
</tr>
<tr>
<td>IZVV</td>
<td></td>
</tr>
</tbody>
</table>
Tools for Data Gathering and Handling

Test planning
- WebScE
  - Test script & route definition
- WebScE
  - Vehicle & driver assignment
- Fleet management

Test execution
- WebScE
  - Test execution
- Test control
- Driver survey

Data handling
- Logstation
- Data transfer chain
- WebScE
  - Instant validation

Database
(Metadata + Log data)
Planning: Vehicle & Driver Assignment

Horst Rechner: Tools for Data Gathering and Handling & Best Practices
Execution: Test Time Information

Definitions

Test planning
- WebSe
  - Test script & route definition
- WebScE
  - Test execution

Test execution
- WebScE
  - Test execution
- Driver survey
  - Fleet management
  - Vehicle & Driver assignment

Data handling
- Logstation
- Data transfer chain
- WebScE
  - Instant validation

Horst Rechner: Tools for Data Gathering and Handling & Best Practices
Data Handling: Logstation

Datenverteilung

- VIS: USB sticks
- RIS: Cellular + Fiber
- Logstation
- Data receive server
- Ringbuffer
- Instant validation evaluators

simTDLogging VIS + RIS + CIS
2012-07-01 until 2012-12-31:
compressed ca. 3 TB
uncompressed ca. 30 TB

This graph covers the first 2 months
Data Handling: Data Transfer Chain

- Logstation
- WebScE
- Test execution
- Data handling
  - Test planning
    - WebScE
    - Test script & route definition
  - WebScE
  - Vehicle & Driver assignment
  - Test control
  - Driver survey
  - Fleet management
  - Instant validation

Diagram showing the data transfer chain with steps:
1. Other data
2. VIS
3. RIS
4. CIS
5. USB hub
6. Data receive server
7. FTP server
8. Ringbuffer
   - (6 TB -> temporary storage)
9. File server
10. Primary evaluators
11. Secondary evaluators
12. USB HDS
13. Transfer log

Arrows indicate the flow of data between these components.
Data Handling: Instant Validation

**Test planning**
- WebScE
- Test script & route definition

**Test execution**
- WebScE
- Test control
- Driver survey

**Data handling**
- Logstation
- Data transfer chain

**Data + Definitions**

Horst Rechner: Tools for Data Gathering and Handling & Best Practices
Best practices

Automation

• **Automate early** (e.g. data transfer chain) for testing and allow partners to ease into setup

• **Automate primary validation** for quality assurance – even if the rulesets and indicators are simple – almost trivial (e.g. datasets vs. vehicles)

• Add **automated safeguards** that inform you when any part of your transfer and validation chain fails

Data quality

• **Check data quality early on!** Check manually – humans can see strange behaviour much better!

• **Automatically create visualizations** of incoming data that can be efficiently parsed by humans
Thank you for your attention.

Any questions?